CSIAM Trans. Appl. Math.
doi: 10.4208/csiam-am.SO-2024-0020

Vol. **x**, No. x, pp. 1-26
xx 2025

Construction and Analysis for Adams Explicit Discretization of High-Index Saddle Dynamics

Shuai Miao[1], Ziqi Liu[2], Lei Zhang[3,*], Pingwen Zhang[1,4] and Xiangcheng Zheng[5]

[1] *School of Mathematical Sciences, Peking University, Beijing 100871, China.*
[2] *Department of Mathematical Sciences, Tsinghua University, Beijing 100084, China.*
[3] *Beijing International Center for Mathematical Research, Center for Quantitative Biology, Center for Machine Learning Research, Peking University, Beijing 100871, China.*
[4] *School of Mathematics and Statistics, Wuhan University, Wuhan 430072, China.*
[5] *School of Mathematics, State Key Laboratory of Cryptography and Digital Economy Security, Shandong University, Jinan, 250100, China.*

Received 3 June 2024; Accepted 21 February 2025

Abstract. Saddle points largely exist in complex systems and play important roles in various scientific problems. High-index saddle dynamics (HiSD) is an efficient method for computing any-index saddle points and constructing solution landscape. In this paper, we propose a two-step Adams explicit scheme for HiSD and analyze its error estimate versus time step. Through careful argumentation and overcoming the difficulties caused by nonlinear coupling and orthogonalisation, we prove that the two-step Adams explicit scheme has second-order accuracy. The theoretical results are further verified by two numerical experiments.

AMS subject classifications: 37M05, 37N30, 65L20

Key words: Saddle point, saddle dynamics, solution landscape, Adams scheme, error estimate.

1 Introduction

Finding the stationary point of the nonlinear multivariate energy function $E(x)$ has been an important concern in many fields of science over the last decades. It has special significance and a wide range of applications in physics, chemistry, and biology, such as critical

*Corresponding author. *Email addresses:* liu-zq21@mails.tsinghua.edu.cn (Z. Liu), pzhang@pku.edu.cn (P. Zhang), miaoshuai@math.pku.edu.cn (S. Miao), zhangl@math.pku.edu.cn (L. Zhang), xzheng@sdu.edu.cn (X. Zheng)

nuclei in phase transitions [10, 38], molecular clusters [1, 6], protein folding [23, 27], artificial neural networks [8], and soft matter [5, 15]. For example, an index-1 saddle point is called a transition state, which means that at a stationary point, the Hessian matrix has and only has one negative eigenvalue. According to Morse theory [25], the Morse index of a non-degenerate saddle point is the maximum dimension of the negative definite subspace of its Hessian matrix $\nabla^2 E(x)$, i.e. the number of negative eigenvalues of the Hessian matrix $\nabla^2 E(x)$.

How to compute the saddle point of $E(x)$ is a very challenging problem [21]. Because of the clear meaning and wide application of index-1 saddle point, extensive methods have been developed in recent decades to find index-1 saddle point. One popular approach is the path-finding method, such as the nudged elastic band method [19, 28] and the string method [7, 9]. The other type method is the surface-walking method, such as the gentlest ascent dynamics (GAD) [11], the dimer-type methods [18, 37, 39], and the activation-relaxation technique [4], etc. In 2013, [2] extended the GAD method to find high-index saddle points. More comprehensive overview can be found in [10, 17, 40]. In 2019, high-index saddle dynamics was proposed to compute any-index saddle points [35], inspired by the optimization-based shrinking dimer method [39]. The HiSD method presents an effective tool for constructing the solution landscape and describes a pathway map starting with a parent state, the highest-index saddle point, that relates to the low-index saddle points and all the minimisers [33]. The solution landscape approach has been applied to investigate several physical systems [16, 29, 30, 32, 36]. In addition, the HiSD method is not limited to the gradient systems, and has been extended to the non-gradient systems by using a generalized high-index saddle dynamics (GHiSD), enabling the calculation of any-index saddle points and solution landscapes of non-gradient systems [34].

Theoretical analyses are essential to improve the confidence of the algorithms. In recent years some scholars have developed some theoretical analyses of saddle point search algorithms. For example, the asymptotic stability and convergence rate of the shrinking dimer dynamics in several different time discretization modes are analysed in [37]. The linear stability and error estimation of the dimer method with preconditioners and line search algorithms are analysed in [13]. Recently, the rate of convergence of numerical scheme for HiSD [35] has been analysed in [22] and it was found that the rate of convergence is mainly related to the local curvature around the saddle-points and the accuracy of the eigenvector computation. There are some other numerical analysis results we present here [12, 14, 20].

The above-mentioned works provide asymptotic convergence results of, e.g. $x_n - x_*$, where x_n and x_* refer to the numerical solution at the n-th iteration and the limit (target saddle point) of the scheme, respectively. By contrast, the difference between x_n and $x(t_n)$ provides an objective measure of the error, whereby $x(t_n)$ represents the exact solution of HiSD at step t_n. This approach allows for the assessment of the convergence of numerical solutions to saddle dynamics, and provides valuable physical insights, including the transition pathway [32, 33]. Incorrect computation of the dynamical pathway may

result in the miss of saddle points and an incomplete solution landscape. Hence, it is essential and significant to estimate the errors of $x_n - x(t_n)$ for HiSD. In a recent study by Zhang *et al.* [41], the error estimation of the explicit Euler discrete scheme of HiSD was conducted, which proved the first-order rate of convergence with respect to time for trajectory positions and eigenvectors. In [42], the boundedness of the solution and the dimer error of the shrinking-dimer saddle dynamics were analyzed under the local Lipschitz conditions.

The explicit Euler scheme has only first-order accuracy, and the multistep method is a good candidate if high accuracy is needed. In this paper we construct the two-step Adams explicit scheme for HiSD and try to prove that it has second-order accuracy with respect to time. Due to the Stiefel manifold constraint in discrete HiSD, the accuracy proof of Adams scheme is far from the analyses of high-order methods of ordinary differential equations and thus has significant difficulties. Through careful argumentation and overcoming the difficulties caused by nonlinear coupling and orthogonalisation, we prove that the two-step Adams explicit scheme has second-order accuracy, which presents the high-accuracy numerical computation of HiSD and provides its numerical analyses for the first time.

The rest of the paper is organised as follows. In Section 2, we present the specific form of HiSD, the assumptions used and the lemma. In Section 3, we construct the two-step Adams explicit scheme for index-1 saddle dynamics and analyse its error with respect to time. In Section 4, we generalise the construction of the scheme and the error estimates to HiSD. In Section 5, we verify our theoretical findings with several numerical experiments. In the last section we generalise the relevant results to GHiSD and draw conclusions.

2 HiSD and assumption

Given a twice Fréchet differentiable energy functional $E(x)$ defined on a real Hilbert space and define the corresponding natural force $F(x)=-\nabla E(x)$ and the negative Hessian $J(x)=-\nabla^2 E(x)$. It is clear that $J(x)=J(x)^\top$. Then the saddle dynamics for an index-k saddle point (k-SD) of $E(x)$ with $1\le k\in\mathbb{N}$ reads [35]

$$\begin{cases} \dfrac{dx}{dt}=\beta\left(I-2\displaystyle\sum_{j=1}^{k}v_jv_j^\top\right)F(x), \\ \dfrac{dv_i}{dt}=\gamma\left(I-v_iv_i^\top-2\displaystyle\sum_{j=1}^{i-1}v_jv_j^\top\right)J(x)v_i, \quad 1\le i\le k, \end{cases} \tag{2.1}$$

where x represents a position variable, v_i (i=1,…,k) are k directional variables, and $\beta,\gamma>0$ are relaxation parameters. Additionally, [35] demonstrates that if the initial values of $\{v_i(t)\}_{i=1}^k$ for (2.1) are unit orthonormal vectors, then $\{v_i(t)\}_{i=1}^k$ are unit orthonormal

vectors for any $t>0$. In particular, the index-1 saddle point dynamics (1-SD) takes the following form, which is equivalent to the gentlest ascent dynamics [11]:

$$\begin{cases} \dfrac{dx}{dt}=\beta\big(I-2vv^\top\big)F(x), \\ \dfrac{dv}{dt}=\gamma\big(I-vv^\top\big)J(x)v. \end{cases} \tag{2.2}$$

Throughout the paper we make the following regular assumptions on the force and the Hessian:

Assumption 2.1. There exists a constant $L>0$ such that the following linearly growth and Lipschitz conditions hold under the standard l^2 norm $\|\cdot\|$ of a vector or a matrix:

$$\begin{aligned} &\|J(x_2)-J(x_1)\|+\|F(x_2)-F(x_1)\|\le L\|x_2-x_1\|, \\ &\|F(x)\|\le L(1+\|x\|), \quad x,x_1,x_2\in\mathbb{R}^N. \end{aligned}$$

Based on Assumption 2.1, [41] demonstrates that the norm of $x(t)$ can be bounded by $\sqrt{Q_T}$ for $t\in[0,T]$, where T is the terminal time, and we thus assume

$$J_T:=\max_{\|x\|\le\sqrt{Q_T}}\|J(x)\|.$$

We cite the discrete Gronwall inequality, which will be frequently used throughout this paper [3].

Lemma 2.1 (Discrete Gronwall Inequality)**.** *Assume that the non-negative sequences* $\{z_n\}_{n\ge1}$ *and* $\{k_n\}_{n\ge1}$ *satisfy*

$$z_n\le\rho+\sum_{j=1}^{n-1}k_jz_j$$

for $n\ge1$ *for some* $\rho\ge0$*. Then*

$$z_n\le\rho\exp\left(\sum_{j=1}^{n-1}k_j\right),\quad n\ge1.$$

In the rest of the work, we use Q to denote a generic positive constant that may assume different values at different occurrences, and use Q_i, $\bar{Q}$ or $\tilde{Q}$ to denote fixed parameters. All constants are independent from time discretization parameters.

3 Numerical analysis for index-1 saddle dynamics

To make the main idea of the proof clearer, in this section we consider the Adams explicit scheme of the index-1 saddle point (2.2) on the time interval $[0,T]$ for some $T>0$.

3.1 Adams explicit scheme

We divide the interval $[0,T]$ into $N_T \in \mathbb{N}$ equal parts, $t_n = n\tau$ for $0 \le n \le N_T$ is called a node, where $\tau = T/N_T$ is the time step size. For first-order ordinary differential equation

$$\frac{\partial y}{\partial t} = f(t,y), \tag{3.1}$$

we calculate (3.1) by the explicit Euler scheme at t_1 as follows:

$$y(t_1) = y(t_0) + \tau f\big(t_0, y(t_0)\big) + T_1^y,$$

and calculate by the second-order Adams explicit scheme at t_n $(2 \le n \le N_T)$ as follows:

$$y(t_n) = y(t_{n-1}) + \frac{3}{2}\tau f\big(t_{n-1}, y(t_{n-1})\big) - \frac{1}{2}\tau f\big(t_{n-2}, y(t_{n-2})\big) + T_n^y,$$

where y refers to x or v, and the truncation error $\|T_1^y\| = \mathcal{O}(\tau^2)$ and $\|T_n^y\| = \mathcal{O}(\tau^3)$ $(2 \le n \le N_T)$. Applying above discretization into (2.2) yields

$$\begin{cases} x(t_1) = x(t_0) + \tau\beta\big(I - 2v(t_0)v(t_0)^\top\big)F\big(x(t_0)\big) + T_1^x, & \text{(3.2a)} \\ v(t_1) = v(t_0) + \tau\gamma\big(I - v(t_0)v(t_0)^\top\big)J\big(x(t_0)\big)v(t_0) + T_1^v, & \text{(3.2b)} \\ x(t_n) = x(t_{n-1}) + \frac{3}{2}\tau\beta\big(I - 2v(t_{n-1})v(t_{n-1})^\top\big)F\big(x(t_{n-1})\big) \\ \qquad\quad - \frac{1}{2}\tau\beta\big(I - 2v(t_{n-2})v(t_{n-2})^\top\big)F\big(x(t_{n-2})\big) + T_n^x, & \text{(3.2c)} \\ v(t_n) = v(t_{n-1}) + \frac{3}{2}\tau\gamma\big(I - v(t_{n-1})v(t_{n-1})^\top\big)J\big(x(t_{n-1})\big)v(t_{n-1}) \\ \qquad\quad - \frac{1}{2}\tau\gamma\big(I - v(t_{n-2})v(t_{n-2})^\top\big)J\big(x(t_{n-2})\big)v(t_{n-2}) + T_n^v. & \text{(3.2d)} \end{cases}$$

Then we drop the truncation errors of (3.2) to obtain the Adams explicit scheme of (2.2) with the approximations $\{x_n, v_n\}_{n=1}^{N_T}$ to $\{x(t_n), v(t_n)\}_{n=1}^{N_T}$

$$\begin{cases} x_1 = x_0 + \tau\beta\big(I - 2v_0v_0^\top\big)F(x_0), & \text{(3.3a)} \\ \tilde{v}_1 = v_0 + \tau\gamma\big(I - v_0v_0^\top\big)J(x_0)v_0, & \text{(3.3b)} \\ v_1 = \dfrac{\tilde{v}_1}{\|\tilde{v}_1\|}, & \text{(3.3c)} \\ x_n = x_{n-1} + \frac{3}{2}\tau\beta\big(I - 2v_{n-1}v_{n-1}^\top\big)F(x_{n-1}) - \frac{1}{2}\tau\beta\big(I - 2v_{n-2}v_{n-2}^\top\big)F(x_{n-2}), & \text{(3.3d)} \\ \tilde{v}_n = v_{n-1} + \frac{3}{2}\tau\gamma\big(I - v_{n-1}v_{n-1}^\top\big)J(x_{n-1})v_{n-1} - \frac{1}{2}\tau\gamma\big(I - v_{n-2}v_{n-2}^\top\big)J(x_{n-2})v_{n-2}, & \text{(3.3e)} \\ v_n = \dfrac{\tilde{v}_n}{\|\tilde{v}_n\|}, & \text{(3.3f)} \end{cases}$$

and the initial conditions $x(0)=x_0, v(0)=v_0, \|v_0\|_2=1$. The Eqs. (3.3a) and (3.3b) are the explicit Euler scheme for calculating x_1 and $\tilde{v}_1$, while the Eqs. (3.3d) and (3.3e) are the Adams explicit scheme for calculating x_n and $\tilde{v}_n$ for $2\leq n\leq N_T$. Since $\tilde{v}_n$ $(1\leq n\leq N_T)$ may not be a unit vector in the sense of the l_2 norm due to the presence of discrete errors, the Eqs. (3.3c) and (3.3f) are used to ensure that vectors v_n $(1\leq n\leq N_T)$ are unit vectors. Notice that $I-2v_nv_n^\top$ $(0\leq n\leq N_T)$ is a Householder matrix, so we get

$$\left\| I-2v_nv_n^\top \right\| =1. \tag{3.4}$$

Based on Assumption 2.1, it is proved in [41] that $\|x(t)\|$ is bounded for $t\in[0,T]$, and then we prove a similar property for $\|x_n\|$ $(1\leq n\leq N_T)$ in (3.3). We thus multiply $x_n^\top$ on both sides of the Eq. (3.3d) and use (3.4) to obtain

$$\begin{aligned}\|x_n\|^2 \leq \|x_n\|\,\|x_{n-1}\| &+\frac{3}{2}\tau\beta\|x_n\|\left\|I-2v_{n-1}v_{n-1}^\top\right\|\|F(x_{n-1})\|\\ &+\frac{1}{2}\tau\beta\|x_n\|\left\|I-2v_{n-2}v_{n-2}^\top\right\|\|F(x_{n-2})\|,\end{aligned}$$

which leads to

$$\begin{aligned}\|x_n\|-\|x_{n-1}\| &\leq \frac{3}{2}\tau\beta\|F(x_{n-1})\|+\frac{1}{2}\tau\beta\|F(x_{n-2})\|\\ &\leq \frac{3}{2}\tau\beta L(1+\|x_{n-1}\|)+\frac{1}{2}\tau\beta L(1+\|x_{n-2}\|),\quad n\geq 2.\end{aligned} \tag{3.5}$$

Similarly, we multiply $x_1^\top$ on both sides of the Eq. (3.3a) to obtain

$$\|x_1\|-\|x_0\|\leq \tau\beta\left\|I-2v_0v_0^\top\right\|\|F(x_0)\|\leq \tau\beta L(1+\|x_0\|).$$

By summing up the inequalities (3.5) from $n=2$ to $n^*\leq N_T$ and adding the above inequality yields

$$\begin{aligned}\|x_{n^*}\|-\|x_0\| &\leq \tau\beta L(1+\|x_0\|)+\frac{3}{2}\tau\beta L\sum_{n=2}^{n^*}(1+\|x_{n-1}\|)+\frac{1}{2}\tau\beta L\sum_{n=2}^{n^*}(1+\|x_{n-2}\|)\\ &\leq \frac{3}{2}\tau\beta L\sum_{n=1}^{n^*}(1+\|x_{n-1}\|)+\frac{1}{2}\tau\beta L\sum_{n=1}^{n^*}(1+\|x_{n-1}\|)\\ &\leq 2T\beta L+2\tau\beta L\sum_{n=1}^{n^*}\|x_{n-1}\|.\end{aligned}$$

Then we apply the discrete Gronwall inequality to conclude that there exists a constant $\bar{Q}_T$ such that $\|x_n\|\leq\bar{Q}_T$ for $0\leq n\leq N_T$, and we thus assume

$$\bar{J}_T:=\max_{\|x\|\leq\bar{Q}_T}\|J(x)\|. \tag{3.6}$$

3.2 Auxiliary estimate of $\tilde{v}_n - v_n$

Inspired by [41], we split $v(t_n)-\tilde{v}_n$ into

$$\big(v(t_n)-v_n\big)+(v_n-\tilde{v}_n)=e_n^v+(v_n-\tilde{v}_n),$$

which introduces an additional error $v_n-\tilde{v}_n$. In particular, we need to show that the $\|v_n-\tilde{v}_n\|$ has the magnitude of $\mathcal{O}(\tau^3)$ in order to preserve the $\mathcal{O}(\tau^2)$ accuracy of the numerical scheme (3.3), which motivates the following estimate.

Lemma 3.1. *Under Assumption* 2.1*, the following estimate holds for τ small enough:*

$$|\|\tilde{v}_n\|-1|\le|\|\tilde{v}_n\|^2-1|\le Q\tau^3,\quad 2\le n\le N_T.$$

Proof. From the schemes of $\tilde{v}_n$ we have $\|\tilde{v}_n\|\ge 1/2$ for τ small enough. From [41, Lemma 3.1 and Corollary 3.2], we have

$$|\|\tilde{v}_1\|-1|\le Q\tau^2,\quad \|v_1-\tilde{v}_1\|\le Q\tau^2. \tag{3.7}$$

We apply the above equation and the Eq. (3.3b), as well as the boundedness of J, to obtain

$$\|v_1-v_0\|=\|v_1-\tilde{v}_1+\tilde{v}_1-v_0\|\le\|v_1-\tilde{v}_1\|+\|\tilde{v}_1-v_0\|\le Q\tau. \tag{3.8}$$

It follows from the boundedness of $\|x_0\|$ that

$$\|x_1-x_0\|\le\tau\beta\big\|\big(I-2v_0v_0^\top\big)F(x_0)\big\|\le\tau\beta\|F(x_0)\|\le Q\tau. \tag{3.9}$$

Denote $g_n=(I-v_nv_n^\top)J(x_n)v_n$, use $v_0^\top g_0=0$ and combine with the Eq. (3.3b) to obtain

$$\begin{aligned}
v_1^\top g_0&=\big(v_1^\top-v_0^\top\big)g_0=\big(\tilde{v}_1^\top-v_0^\top-\big(\tilde{v}_1^\top-v_1^\top\big)\big)g_0\\
&=\left(\tilde{v}_1^\top-v_0^\top-\left(1-\frac{1}{\|\tilde{v}_1\|}\right)\tilde{v}_1^\top\right)g_0\\
&=\tau\gamma g_0^\top g_0-\left(\frac{\|\tilde{v}_1\|-1}{\|\tilde{v}_1\|}\right)\tilde{v}_1^\top g_0.
\end{aligned}\tag{3.10}$$

Subtract g_0 from g_1 to obtain

$$\begin{aligned}
g_1-g_0&=\big(I-v_1v_1^\top\big)J(x_1)v_1-\big(I-v_0v_0^\top\big)J(x_0)v_0\\
&=J(x_1)v_1-J(x_0)v_0+v_0v_0^\top J(x_0)v_0-v_1v_1^\top J(x_1)v_1\\
&=J(x_1)v_1-J(x_0)v_1+J(x_0)v_1-J(x_0)v_0+v_0v_0^\top J(x_0)v_0\\
&\quad-v_1v_0^\top J(x_0)v_0+v_1v_0^\top J(x_0)v_0-v_1v_1^\top J(x_0)v_0+v_1v_1^\top J(x_0)v_0\\
&\quad-v_1v_1^\top J(x_1)v_0+v_1v_1^\top J(x_1)v_0-v_1v_1^\top J(x_1)v_1.
\end{aligned}$$

We then apply Assumption 2.1, (3.8) and (3.9), as well as the Eq. (3.3a), to find

$$\begin{aligned}\|g_1-g_0\| &\le \|J(x_1)-J(x_0)\|+\|J(x_0)\|\|v_1-v_0\|+\|v_0-v_1\|\|J(x_0)\| \\ &\quad +\|v_0-v_1\|\|J(x_0)\|+\|J(x_0)-J(x_1)\|+\|J(x_1)\|\|v_0-v_1\| \\ &\le 2L\|x_1-x_0\|+4\bar{J}_T\|v_0-v_1\| \le Q\tau. \end{aligned} \tag{3.11}$$

By $\|v_1\|=1$ we directly calculate $\|\tilde{v}_2\|^2$ in (3.3) and apply (3.10) to obtain

$$\begin{aligned}\|\tilde{v}_2\|^2 &= \tilde{v}_2^\top \tilde{v}_2 = \left(v_1+\frac{3}{2}\tau\gamma g_1-\frac{1}{2}\tau\gamma g_0\right)^\top \left(v_1+\frac{3}{2}\tau\gamma g_1-\frac{1}{2}\tau\gamma g_0\right) \\ &= 1+\frac{9}{4}\tau^2\gamma^2 g_1^\top g_1+\frac{1}{4}\tau^2\gamma^2 g_0^\top g_0-\frac{3}{2}\tau^2\gamma^2 g_1^\top g_0-\tau\gamma v_1^\top g_0 \\ &= 1+\frac{9}{4}\tau^2\gamma^2 g_1^\top g_1+\frac{1}{4}\tau^2\gamma^2 g_0^\top g_0-\frac{3}{2}\tau^2\gamma^2 g_1^\top g_0-\tau^2\gamma^2 g_0^\top g_0+\tau\gamma\left(\frac{\|\tilde{v}_1\|-1}{\|\tilde{v}_1\|}\right)\tilde{v}_1^\top g_0 \\ &= 1+\frac{9}{4}\tau^2\gamma^2 g_1^\top g_1-\frac{3}{4}\tau^2\gamma^2 g_0^\top g_0-\frac{3}{2}\tau^2\gamma^2 g_1^\top g_0+\tau\gamma\left(\frac{\|\tilde{v}_1\|-1}{\|\tilde{v}_1\|}\right)\tilde{v}_1^\top g_0 \\ &= 1+\frac{3}{4}\tau^2\gamma^2\left(3g_1^\top(g_1-g_0)+g_0^\top(g_1-g_0)\right)+\tau\gamma\left(\frac{\|\tilde{v}_1\|-1}{\|\tilde{v}_1\|}\right)\tilde{v}_1^\top g_0. \end{aligned} \tag{3.12}$$

Then we use (3.7), (3.11) and (3.12) to obtain

$$\begin{aligned}\left|\|\tilde{v}_2\|^2-1\right| &\le \frac{3}{4}\tau^2\gamma^2\left(3\left\|g_1^\top\right\|\|g_1-g_0\|+\left\|g_0^\top\right\|\|g_1-g_0\|\right) \\ &\quad +\tau\gamma\frac{|\|\tilde{v}_1\|-1|}{\|\tilde{v}_1\|}\left\|\tilde{v}_1^\top\right\|\|g_0\| \le Q\tau^3. \end{aligned}$$

We incorporate this estimate with

$$|\|\tilde{v}_2\|-1| \le |(\|\tilde{v}_2\|+1)(\|\tilde{v}_2\|-1)| = |\|\tilde{v}_2\|^2-1| \le Q\tau^3 \tag{3.13}$$

and

$$\|\tilde{v}_n-v_n\| = \frac{\|\tilde{v}_n\|}{\|\tilde{v}_n\|}|\|\tilde{v}_n\|-1| = |\|\tilde{v}_n\|-1| \tag{3.14}$$

to get the statement of this lemma with $n=2$.

Following a derivation similar to (3.11), we obtain for $m\ge 3$

$$\|g_{m-1}-g_{m-2}\| \le Q\|x_{m-1}-x_{m-2}\|+Q\|v_{m-1}-v_{m-2}\|. \tag{3.15}$$

Following a derivation similar to (3.9), we obtain

$$\|x_{m-1}-x_{m-2}\| \le Q\tau. \tag{3.16}$$

We apply the Eq. (3.3e), as well as the boundedness of J, to obtain

$$\|\tilde{v}_{m-1}-v_{m-2}\| \le Q\tau. \tag{3.17}$$

Then we use (3.15)-(3.17) to obtain

$$\begin{aligned}\|g_{m-1}-g_{m-2}\| &\le Q\|x_{m-1}-x_{m-2}\|+Q\|v_{m-1}-v_{m-2}\|\\ &\le Q\tau+Q(\|v_{m-1}-\tilde{v}_{m-1}\|+\|\tilde{v}_{m-1}-v_{m-2}\|)\\ &\le Q\tau+Q\|v_{m-1}-\tilde{v}_{m-1}\|. \end{aligned}\tag{3.18}$$

By $v_{m-2}^\top g_{m-2}=0$ we invoke the Eq. (3.3e) to obtain

$$\left|\tilde{v}_{m-1}^\top g_{m-2}\right|\le Q\tau,\tag{3.19}$$

where $m\ge 3$. By $\|v_{m-1}\|=1$ we then multiply $\tilde{v}_m^\top$ on both sides of the Eq. (3.3b) to obtain

$$\begin{aligned}\|\tilde{v}_m\|^2 &= \tilde{v}_m^\top\tilde{v}_m\\ &= \left(v_{m-1}+\frac{3}{2}\tau\gamma g_{m-1}-\frac{1}{2}\tau\gamma g_{m-2}\right)^\top\left(v_{m-1}+\frac{3}{2}\tau\gamma g_{m-1}-\frac{1}{2}\tau\gamma g_{m-2}\right)\\ &= 1+\frac{9}{4}\tau^2\gamma^2 g_{m-1}^\top g_{m-1}+\frac{1}{4}\tau^2\gamma^2 g_{m-2}^\top g_{m-2}-\frac{3}{2}\tau^2\gamma^2 g_{m-1}^\top g_{m-2}\\ &\quad -\tau\gamma v_{m-1}^\top g_{m-2}\\ &= 1+\frac{9}{4}\tau^2\gamma^2 g_{m-1}^\top g_{m-1}+\frac{1}{4}\tau^2\gamma^2 g_{m-2}^\top g_{m-2}-\frac{3}{2}\tau^2\gamma^2 g_{m-1}^\top g_{m-2}\\ &\quad -\frac{3}{2}\tau^2\gamma^2 g_{m-2}^\top g_{m-2}+\frac{1}{2}\tau^2\gamma^2 g_{m-3}^\top g_{m-2}+\tau\gamma\left(\frac{\|\tilde{v}_{m-1}\|-1}{\|\tilde{v}_{m-1}\|}\right)\tilde{v}_{m-1}^\top g_{m-2}\\ &= 1+\frac{1}{4}\tau^2\gamma^2(9g_{m-1}^\top(g_{m-1}-g_{m-2})+3(g_{m-1}-g_{m-2})^\top g_{m-2}\\ &\quad -2(g_{m-2}-g_{m-3})^\top g_{m-2})+\tau\gamma\left(\frac{\|\tilde{v}_{m-1}\|-1}{\|\tilde{v}_{m-1}\|}\right)\tilde{v}_{m-1}^\top g_{m-2}. \end{aligned}$$

Applying (3.19) yields

$$\begin{aligned}|\|\tilde{v}_m\|-1| &\le |(\|\tilde{v}_m\|-1)(\|\tilde{v}_m\|+1)| = |\|\tilde{v}_m\|^2-1|\\ &= \frac{1}{4}\tau^2\gamma^2\left[9\|g_{m-1}^\top\|\|g_{m-1}-g_{m-2}\|+3\|g_{m-1}-g_{m-2}\|\|g_{m-2}\|\right.\\ &\qquad \left.+2\|g_{m-2}-g_{m-3}\|\|g_{m-2}\|\right]+\tau\gamma\frac{|\|\tilde{v}_{m-1}\|-1|}{\|\tilde{v}_{m-1}\|}|\tilde{v}_{m-1}^\top g_{m-2}|\\ &\le Q\tau^2\|v_{m-1}-\tilde{v}_{m-1}\|+Q\tau^2\|v_{m-2}-\tilde{v}_{m-2}\|+Q\tau^3+Q\tau^2|\|\tilde{v}_{m-1}\|-1|. \end{aligned}$$

Combining the above equations, we obtain

$$|\|\tilde{v}_m\|-1|\le Q\tau^2|\|\tilde{v}_{m-1}\|-1|+Q\tau^2|\|\tilde{v}_{m-2}\|-1|+Q\tau^3.\tag{3.20}$$

Let

$$a=\frac{Q\tau^2+\sqrt{Q^2\tau^4+4Q\tau^2}}{2},\quad b=\frac{-Q\tau^2+\sqrt{Q^2\tau^4+4Q\tau^2}}{2} \tag{3.21}$$

such that $a-b=Q\tau^2$ and $ab=Q\tau^2$. Therefore, (3.20) can be rewritten as

$$|\|\tilde{v}_m\|-1|+b|\|\tilde{v}_{m-1}\|-1|\leq a(|\|\tilde{v}_{m-1}\|-1|+b|\|\tilde{v}_{m-2}\|-1|)+Q\tau^3. \tag{3.22}$$

Let

$$p_m=|\|\tilde{v}_m\|-1|+b|\|\tilde{v}_{m-1}\|-1|$$

such that the above inequality can be further rewritten as

$$p_m\leq ap_{m-1}+Q\tau^3.$$

We repeatedly apply this relation to get

$$p_m\leq a^{m-2}p_2+Q\tau^3\sum_{i=0}^{m-3}a^i.$$

From (3.7), (3.13) and (3.21), which implies $b\leq Q\tau$, we have

$$|p_2|\leq|\|\tilde{v}_2\|-1|+b|\|\tilde{v}_1\|-1|\leq Q\tau^3.$$

Then for τ small enough such that $a<1$, we have

$$|\|\tilde{v}_m\|-1|\leq p_m\leq a^{m-2}p_2+Q\tau^3\sum_{i=0}^{m-3}a^i\leq Q\tau^3,$$

which, together with (3.14), completes the proof. □

3.3 Error estimates

Define the errors by $e_n^x:=x(t_n)-x_n$ and $e_n^v:=v(t_n)-v_n$ for $1\leq n\leq N_T$. We first bound e_n^x in terms of e_n^v in the following theorem.

Theorem 3.1. *Suppose Assumption* 2.1 *holds. Then the following estimate holds:*

$$\|e_n^x\|\leq Q\tau\sum_{m=1}^{n-1}\|e_m^v\|+Q\tau^2,\quad 1\leq n\leq N_T.$$

Here Q depends on L,T and β but is independent from τ,n and N_T.

Proof. We subtract the Eq. (3.2a) from that of (3.3) to obtain

$$\|e_1^x\| \le \|T_1^x\|. \tag{3.23}$$

We subtract the Eq. (3.2a) from the Eq. (3.3d) to obtain

$$\begin{aligned}
e_n^x =& e_{n-1}^x + \frac{3}{2}\tau\beta\big(F\big(x(t_{n-1})\big) - F(x_{n-1})\big) \\
&-3\tau\beta\big[v(t_{n-1})v(t_{n-1})^\top F\big(x(t_{n-1})\big) - v_{n-1}v_{n-1}^\top F(x_{n-1})\big] \\
&+\frac{1}{2}\tau\beta\big(F(x_{n-2}) - F\big(x(t_{n-2})\big)\big) \\
&-\tau\beta\big[v(t_{n-2})v(t_{n-2})^\top F\big(x(t_{n-2})\big) - v_{n-2}v_{n-2}^\top F(x_{n-2})\big] + T_n^x \\
=& e_{n-1}^x + \frac{3}{2}\tau\beta\big(F\big(x(t_{n-1})\big) - F(x_{n-1})\big) \\
&-3\tau\beta\big[e_{n-1}^v v(t_{n-1})^\top F\big(x(t_{n-1})\big) + v_{n-1}(e_{n-1}^v)^\top F\big(x(t_{n-1})\big) \\
&\qquad + v_{n-1}v_{n-1}^\top\big(F\big(x(t_{n-1})\big) - F(x_{n-1})\big)\big] \\
&+\frac{1}{2}\tau\beta\big(F(x_{n-2}) - F\big(x(t_{n-2})\big)\big) \\
&-\tau\beta\big[e_{n-2}^v v(t_{n-2})^\top F\big(x(t_{n-2})\big) + v_{n-2}\big(e_{n-2}^v\big)^\top F\big(x(t_{n-2})\big) \\
&\qquad + v_{n-2}v_{n-2}^\top\big(F\big(x(t_{n-2})\big) - F(x_{n-2})\big)\big] + T_n^x, \quad n \ge 2.
\end{aligned}$$

We then apply the Assumption 2.1 to find

$$\begin{aligned}
\|e_n^x\| \le& \|e_{n-1}^x\| + \frac{3}{2}\tau\beta\big\|F\big(x(t_{n-1})\big) - F(x_{n-1})\big\| \\
&+3\tau\beta\big[\big\|e_{n-1}^v\big\|\|v(t_{n-1})\|\big\|F\big(x(t_{n-1})\big)\big\| + \|v_{n-1}\|\big\|e_{n-1}^v\big\|\big\|F\big(x(t_{n-1})\big)\big\| \\
&\qquad + \|v_{n-1}\|\|v_{n-1}\|\big\|F\big(x(t_{n-1})\big) - F(x_{n-1})\big\|\big] \\
&+\frac{1}{2}\tau\beta\big\|F\big(x(t_{n-2})\big) - F(x_{n-2})\big\| \\
&+\tau\beta\big[\big\|e_{n-2}^v\big\|\|v(t_{n-2})\|\big\|F\big(x(t_{n-2})\big)\big\| + \|v_{n-2}\|\big\|e_{n-2}^v\big\|\big\|F\big(x(t_{n-2})\big)\big\| \\
&\qquad + \|v_{n-2}\|\|v_{n-2}\|\big\|F\big(x(t_{n-2})\big) - F(x_{n-2})\big\|\big] + \|T_n^x\| \\
\le& \|e_{n-1}^x\| + \frac{3}{2}\tau\beta L\|e_{n-1}^x\| \\
&+3\tau\beta\big[\big\|e_{n-1}^v\big\|L(1+\sqrt{Q_T}) + \|e_{n-1}^v\|L(1+\sqrt{Q_T}) + L\|e_{n-1}^x\|\big] + \frac{1}{2}\tau\beta L\|e_{n-2}^x\| \\
&+\tau\beta\big[\big\|e_{n-2}^v\big\|L(1+\sqrt{Q_T}) + \big\|e_{n-2}^v\big\|L(1+\sqrt{Q_T}) + L\|e_{n-2}^x\|\big] + \|T_n^x\| \\
\le& \|e_{n-1}^x\| + \frac{3}{2}\tau\beta L\|e_{n-1}^x\| + 6\tau\beta L(1+\sqrt{Q_T})\|e_{n-1}^v\| \\
&+\frac{3}{2}\tau\beta L\|e_{n-2}^x\| + 2\tau\beta L(1+\sqrt{Q_T})\|e_{n-2}^v\| + \|T_n^x\|.
\end{aligned}$$

Adding this equation from $n=2$ to n_* and using (3.23) leads to

$$\|e_{n_*}^x\| \le 3\tau\beta L \sum_{n=1}^{n_*} \|e_{n-1}^x\| + \sum_{n=1}^{n_*} \left[8\tau\beta L(1+\sqrt{Q_T})\|e_{n-1}^v\| + \|T_n^x\|\right] + \|T_1^x\|.$$

Then an application of the discrete Gronwall inequality in Lemma 2.1 yields

$$\|e_{n_*}^x\| \le Q\left[\sum_{n=1}^{n_*} \left(\tau\|e_{n-1}^v\| + \|T_n^x\|\right) + \|T_1^x\|\right].$$

We incorporate this estimate with $\|T_n^x\| = \mathcal{O}(\tau^3)$, $\|T_1^x\| = \mathcal{O}(\tau^2)$ to complete the proof. □

Based on the derived results, we prove the error estimates of the Adams explicit scheme (3.3) in the following theorem.

Theorem 3.2. *Under Assumption* 2.1*, the following estimate holds:*

$$\|e_n^x\| + \|e_n^v\| \le Q\tau^2, \quad 1 \le n \le N_T.$$

Here Q depends on L, T and β but is independent from τ, n and N_T.

Proof. We subtract the Eq. (3.2d) from the Eq. (3.3e) to obtain

$$\begin{aligned}
v(t_n) - \tilde{v}_n &= e_{n-1}^v + \frac{3}{2}\tau\gamma\big(J\big(x(t_{n-1})\big)v(t_{n-1}) - J(x_{n-1})v_{n-1}\big) \\
&\quad - \frac{3}{2}\tau\gamma\big[v(t_{n-1})v(t_{n-1})^\top J\big(x(t_{n-1})\big)v(t_{n-1}) - v_{n-1}v_{n-1}^\top J(x_{n-1})v_{n-1}\big] \\
&\quad - \frac{1}{2}\tau\gamma\big(J\big(x(t_{n-2})\big)v(t_{n-2}) - J(x_{n-2})v_{n-2}\big) \\
&\quad + \frac{1}{2}\tau\gamma\big[v(t_{n-2})v(t_{n-2})^\top J\big(x(t_{n-2})\big)v(t_{n-2}) - v_{n-2}v_{n-2}^\top J(x_{n-2})v_{n-2}\big] + T_n^v \\
&= e_{n-1}^v + \frac{3}{2}\tau\gamma\big[\big(J\big(x(t_{n-1})\big) - J(x_{n-1})\big)v(t_{n-1}) + J(x_{n-1})e_{n-1}^v\big] \\
&\quad - \frac{3}{2}\tau\gamma\big[e_{n-1}^v v(t_{n-1})^\top J\big(x(t_{n-1})\big)v(t_{n-1}) + v_{n-1}\big(e_{n-1}^v\big)^\top J\big(x(t_{n-1})\big)v(t_{n-1}) \\
&\qquad\quad + v_{n-1}v_{n-1}^\top\big(J\big(x(t_{n-1})\big) - J(x_{n-1})\big)v(t_{n-1}) + v_{n-1}v_{n-1}^\top J(x_{n-1})e_{n-1}^v\big] \\
&\quad - \frac{1}{2}\tau\gamma\big[\big(J\big(x(t_{n-2})\big) - J(x_{n-2})\big)v(t_{n-2}) + J(x_{n-2})e_{n-2}^v\big] \\
&\quad + \frac{1}{2}\tau\gamma\big[e_{n-2}^v v(t_{n-2})^\top J\big(x(t_{n-2})\big)v(t_{n-2}) + v_{n-2}\big(e_{n-2}^v\big)^\top J\big(x(t_{n-2})\big)v(t_{n-2}) \\
&\qquad\quad + v_{n-2}v_{n-2}^\top\big(J\big(x(t_{n-2})\big) - J(x_{n-2})\big)v(t_{n-2}) + v_{n-2}v_{n-2}^\top J(x_{n-2})e_{n-2}^v\big] + T_n^v,
\end{aligned}$$

which leads to

$$\begin{aligned}\|v(t_n)-\tilde{v}_n\| \le \|e^v_{n-1}\| &+ \frac{3}{2}\tau\gamma\left[L\|e^x_{n-1}\|+\bar{J}_T\|e^v_{n-1}\|\right]\\ &+\frac{3}{2}\tau\gamma\left[2J_T\|e^v_{n-1}\|+L\|e^x_{n-1}\|+\bar{J}_T\|e^v_{n-1}\|\right]+\frac{1}{2}\tau\gamma\left[L\|e^x_{n-2}\|+\bar{J}_T\|e^v_{n-2}\|\right]\\ &+\frac{1}{2}\tau\gamma\left[2J_T\|e^v_{n-2}\|+L\|e^x_{n-2}\|+\bar{J}_T\|e^v_{n-2}\|\right]+T^v_n. \end{aligned} \tag{3.24}$$

We split $v(t_n)-\tilde{v}_n$ as

$$\big(v(t_n)-v_n\big)+(v_n-\tilde{v}_n)=e^v_n+(v_n-\tilde{v}_n),$$

and apply Lemma 3.1 and Theorem 3.1 for (3.24) to get

$$\begin{aligned}\|e^v_n\| &\le \|e^v_{n-1}\|+\|v_n-\tilde{v}_n\|+Q\tau\big(\|e^x_{n-1}\|+\|e^v_{n-1}\|\big)\\ &\quad+Q\tau\big(\|e^x_{n-2}\|+\|e^v_{n-2}\|\big)+\|T^v_n\|\\ &\le \|e^v_{n-1}\|+Q\tau\|e^v_{n-1}\|+Q\tau\|e^v_{n-2}\|+Q\tau^2\sum_{m=1}^{n-1}\|e^v_m\|+Q\tau^3.\end{aligned}$$

Adding this estimates from $n=2$ to n_* for $2\le n_*\le N_T$ and using

$$\tau^2\sum_{n=1}^{n_*}\sum_{m=1}^{n-1}\|e^v_m\|=\tau^2\sum_{m=1}^{n_*-1}\sum_{n=m+1}^{n_*}\|e^v_m\|\le T\tau\sum_{m=1}^{n_*-1}\|e^v_m\|,$$

we get

$$\|e^v_{n_*}\|\le\|e^v_1\|+Q\tau\sum_{n=1}^{n_*-1}\|e^v_n\|+Q\tau^2.$$

From [41, inequality (3.15)], we have $\|e^v_1\|\le Q\tau^2$. Then an application of the discrete Gronwall inequality leads to $\|e^v_n\|\le Q\tau^2$ for $1\le n\le N_T$. Plugging this estimate back to the conclusion of Theorem 3.1 yields the estimate of $\|e^x_n\|$ and we thus complete the proof. □

4 Numerical analysis for index-k saddle dynamics

We study the Adams explicit approximation of k-SD (2.1) for some $k>1$, which has k eigenvectors and thus needs additional orthogonalization process that significantly complicates the analysis.

4.1 Adams explicit scheme

Similar to Section 3.1, Adams explicit scheme in the continuous case for k-SD (2.1) reads

$$\begin{cases} x(t_1)=x(t_0)+\tau\beta\Big(I-2\sum_{j=1}^{k}v_j(t_0)v_j(t_0)^\top\Big)F\big(x(t_0)\big)+T_1^x,\\ v_i(t_1)=v_i(t_0)+\tau\gamma\Big(I-v_i(t_0)v_i(t_0)^\top-2\sum_{j=1}^{i-1}v_j(t_0)v_j(t_0)^\top\Big)J\big(x(t_0)\big)v_i(t_0)+T_1^{v_i},\\ x(t_n)=x(t_{n-1})+\frac{3}{2}\tau\beta\Big(I-2\sum_{j=1}^{k}v_j(t_{n-1})v_j(t_{n-1})^\top\Big)F\big(x(t_{n-1})\big)\\ \qquad -\frac{1}{2}\tau\beta\Big(I-2\sum_{j=1}^{k}v_j(t_{n-2})v_j(t_{n-2})^\top\Big)F(x(t_{n-2}))+T_n^x,\\ v_i(t_n)=v_i(t_{n-1})\\ \qquad +\frac{3}{2}\tau\gamma\Big(I-v_i(t_{n-1})v_i(t_{n-1})^\top-2\sum_{j=1}^{i-1}v_j(t_{n-1})v_j(t_{n-1})^\top\Big)J\big(x(t_{n-1})\big)v_i(t_{n-1})\\ \qquad -\frac{1}{2}\tau\gamma\Big(I-v_i(t_{n-2})v_i(t_{n-2})^\top-2\sum_{j=1}^{i-1}v_j(t_{n-2})v_j(t_{n-2})^\top\Big)J\big(x(t_{n-2})\big)v_i(t_{n-2})+T_n^{v_i} \end{cases}$$

for $1\le i\le k$. Then we drop the truncation errors to obtain the Adams explicit scheme of (2.1)

$$\begin{cases} x_1=x_0+\tau\beta\Big(I-2\sum_{j=1}^{k}v_{j,0}v_{j,0}^\top\Big)F(x_0), & (4.1a)\\ \tilde{v}_{i,1}=v_{i,0}+\tau\gamma\Big(I-v_{i,0}v_{i,0}^\top-2\sum_{j=1}^{i-1}v_{j,0}v_{j,0}^\top\Big)J(x_0)v_{i,0}, \quad 1\le i\le k, & (4.1b)\\ v_{i,1}=\frac{1}{Y_{i,1}}\Big(\tilde{v}_{i,1}-\sum_{j=1}^{i-1}(\tilde{v}_{i,1}^\top v_{j,1})v_{j,1}\Big), \quad 1\le i\le k, & (4.1c)\\ x_n=x_{n-1}+\frac{3}{2}\tau\beta\Big(I-2\sum_{j=1}^{k}v_{j,n-1}v_{j,n-1}^\top\Big)F(x_{n-1})\\ \qquad -\frac{1}{2}\tau\beta\Big(I-2\sum_{j=1}^{k}v_{j,n-2}v_{j,n-2}^\top\Big)F(x_{n-2}), & (4.1d)\\ \tilde{v}_{i,n}=v_{i,n-1}+\frac{3}{2}\tau\gamma\Big(I-v_{i,n-1}v_{i,n-1}^\top-2\sum_{j=1}^{i-1}v_{j,n-1}v_{j,n-1}^\top\Big)J(x_{n-1})v_{i,n-1}\\ \qquad -\frac{1}{2}\tau\gamma\Big(I-v_{i,n-2}v_{i,n-2}^\top-2\sum_{j=1}^{i-1}v_{j,n-2}v_{j,n-2}^\top\Big)J(x_{n-2})v_{i,n-2}, \quad 1\le i\le k, & (4.1e)\\ v_{i,n}=\frac{1}{Y_{i,n}}\Big(\tilde{v}_{i,n}-\sum_{j=1}^{i-1}(\tilde{v}_{i,n}^\top v_{j,n})v_{j,n}\Big), \quad 1\le i\le k & (4.1f) \end{cases}$$

for $2\leq n\leq N_T$ equipped with the initial conditions

$$x(0)=x_0,\quad v_i(0)=v_{i,0}\quad\text{with}\quad v_{i,0}^\top v_{j,0}=\delta_{i,j},\quad 1\leq i,j\leq k,$$

where $\delta_{i,j}=1$ if $i=j$ and $\delta_{i,j}=0$ otherwise. The Eq. (4.1f) is the standard Gram-Schmidt normalized orthogonalization procedure in order to preserve the orthonormal property of the vectors as in the continuous problem (2.1), and $Y_{i,n}$ stands for the norm of the corresponding vector, i.e.

$$Y_{i,n}:=\left\|\tilde{v}_{i,n}-\sum_{j=1}^{i-1}\left(\tilde{v}_{i,n}^\top v_{j,n}\right)v_{j,n}\right\|=\left(\|\tilde{v}_{i,n}\|^2-\sum_{j=1}^{i-1}\left(\tilde{v}_{i,n}^\top v_{j,n}\right)^2\right)^{\frac{1}{2}}.$$

Similar to the derivation of $\bar{J}_T$ in (3.6), we could conclude that there exists a positive constant $\hat{J}_T$ independent from N_T and τ such that $\max_{1\leq n\leq N_T}\|J(x_n)\|\leq\hat{J}_T$.

4.2 Auxiliary estimates

We first present the following lemma for future use.

Lemma 4.1. *There exists a function $g:\mathcal{R}^+\to\mathcal{R}^+$ such that if a set of vectors $\{w_i\}_{i=1}^k$ satisfy $|w_i^\top w_j-\delta_{i,j}|\leq C\tau^3$ for $1\leq i\leq j\leq k$ for some constant $C>0$ and $\{\tilde{w}_i\}_{i=1}^k$ are orthonormal vectors generated from $\{w_i\}_{i=1}^k$ via the Gram-Schmidt process, the following estimate holds for τ sufficiently small*

$$\|\tilde{w}_i-w_i\|\leq g(C)\tau^3,\quad 1\leq i\leq k.$$

Proof. The proof is similar to the analysis in [41, Lemma 4.2] and is omitted here. □

Lemma 4.2. *Under the Assumption* 2.1*, we have*

$$\|v_{i,n}-\tilde{v}_{i,n}\|\leq Q\tau^3$$

for $1\leq i\leq k$ and $2\leq n\leq N_T$ for the Adams explicit scheme (4.1) *for some constant $Q>0$.*

Proof. It is shown in [41] that

$$\|v_{i,1}-\tilde{v}_{i,1}\|\leq Q_1\tau^2,\quad 1\leq i\leq k,\tag{4.2}$$

for some $Q_1>0$. Denoting

$$g_{i,n}=\left(I-v_{i,n}v_{i,n}^\top-2\sum_{j=1}^{i-1}v_{j,n}v_{j,n}^\top\right)J(x_n)v_{i,n}$$

for $1\leq i\leq k$ and $0\leq n\leq N_T$, based on the symmetry of J, we can easily verify

$$v_{m,n}^T g_{i,n}+v_{i,n}^T g_{m,n}=0\tag{4.3}$$

for $1\leq m,i\leq k$ and $0\leq n\leq N_T$. We directly calculate to get

$$\begin{aligned}
&(v_{i,n-1}-v_{i,n-2})^\top g_{m,n-2}\\
=&(\tilde{v}_{i,n-1}-v_{i,n-2}+v_{i,n-1}-\tilde{v}_{i,n-1})^\top g_{m,n-2}\\
=&\frac{3}{2}\tau\gamma g_{i,n-2}g_{m,n-2}-\frac{1}{2}\tau\gamma g_{i,n-3}g_{m,n-2}+(v_{i,n-1}-\tilde{v}_{i,n-1})^\top g_{m,n-2},\\
&(v_{m,n-1}-v_{m,n-2})^\top g_{i,n-2}\\
=&(\tilde{v}_{m,n-1}-v_{m,n-2}+v_{m,n-1}-\tilde{v}_{m,n-1})^\top g_{i,n-2}\\
=&\frac{3}{2}\tau\gamma g_{m,n-2}g_{i,n-2}-\frac{1}{2}\tau\gamma g_{m,n-3}g_{i,n-2}+(v_{m,n-1}-\tilde{v}_{m,n-1})^\top g_{i,n-2}.
\end{aligned}$$

Using the first equation of (4.1) yields

$$\begin{aligned}
v_{m,1}-v_{m,0}&=\tilde{v}_{m,1}-v_{m,0}+v_{m,1}-\tilde{v}_{m,0}\\
&=\tau\gamma g_{m,0}+v_{m,1}-\tilde{v}_{m,0},
\end{aligned}\tag{4.4}$$

$$\begin{aligned}
v_{i,1}-v_{i,0}&=\tilde{v}_{i,1}-v_{i,0}+v_{i,1}-\tilde{v}_{i,0}\\
&=\tau\gamma g_{i,0}+v_{i,1}-\tilde{v}_{i,0}.
\end{aligned}\tag{4.5}$$

By (4.3), we directly calculate the product $\tilde{v}_{m,2}^\top\tilde{v}_{i,2}$ for $1\leq m,i\leq k$

$$\begin{aligned}
\tilde{v}_{m,2}^\top\tilde{v}_{i,2}=&\left(v_{m,1}+\frac{3}{2}\tau\gamma g_{m,1}-\frac{1}{2}\tau\gamma g_{m,0}\right)^\top\left(v_{i,1}+\frac{3}{2}\tau\gamma g_{i,1}-\frac{1}{2}\tau\gamma g_{i,0}\right)\\
=&\delta_{m,i}+\frac{9}{4}\tau^2\gamma^2 g_{m,1}^\top g_{i,1}-\frac{3}{4}\tau^2\gamma^2 g_{m,1}^\top g_{i,0}-\frac{3}{4}\tau^2\gamma^2 g_{m,0}^\top g_{i,1}\\
&+\frac{1}{4}\tau^2\gamma^2 g_{m,0}^\top g_{i,0}+\frac{3}{2}\tau\gamma v_{m,1}^\top g_{i,1}-\frac{1}{2}\tau\gamma v_{m,1}^\top g_{i,0}\\
&+\frac{3}{2}\tau\gamma v_{i,1}^\top g_{m,1}-\frac{1}{2}\tau\gamma v_{i,1}^\top g_{m,0}.
\end{aligned}$$

Using (4.3), we have

$$\begin{aligned}
\tilde{v}_{m,2}^\top\tilde{v}_{i,2}=&\delta_{m,i}+\frac{9}{4}\tau^2\gamma^2 g_{m,1}^\top g_{i,1}-\frac{3}{4}\tau^2\gamma^2 g_{m,1}^\top g_{i,0}-\frac{3}{4}\tau^2\gamma^2 g_{m,0}^\top g_{i,1}\\
&+\frac{1}{4}\tau^2\gamma^2 g_{m,0}^\top g_{i,0}+\frac{1}{2}\tau\gamma v_{m,0}^\top g_{i,0}-\frac{1}{2}\tau\gamma v_{m,1}^\top g_{i,0}\\
&+\frac{1}{2}\tau\gamma v_{i,0}^\top g_{m,0}-\frac{1}{2}\tau\gamma v_{i,1}^\top g_{m,0}\\
=&\delta_{m,i}+\frac{9}{4}\tau^2\gamma^2 g_{m,1}^\top g_{i,1}-\frac{3}{4}\tau^2\gamma^2 g_{m,1}^\top g_{i,0}-\frac{3}{4}\tau^2\gamma^2 g_{m,0}^\top g_{i,1}\\
&+\frac{1}{4}\tau^2\gamma^2 g_{m,0}^\top g_{i,0}-\frac{1}{2}\tau\gamma(v_{m,1}-v_{m,0})^\top g_{i,0}-\frac{1}{2}\tau\gamma(v_{i,1}-v_{i,0})^\top g_{m,0}.
\end{aligned}$$

Substituting (4.4) and (4.5) into the above equation and by factorisation we get

$$\begin{aligned}\tilde{v}_{m,2}^{\top}\tilde{v}_{i,2}=&\delta_{m,i}+\frac{9}{4}\tau^2\gamma^2 g_{m,1}^{\top}g_{i,1}-\frac{3}{4}\tau^2\gamma^2 g_{m,1}^{\top}g_{i,0}-\frac{3}{4}\tau^2\gamma^2 g_{m,0}^{\top}g_{i,1}\\&-\frac{3}{4}\tau^2\gamma^2 g_{m,0}^{\top}g_{i,0}-\frac{1}{2}\tau^2\gamma^2(v_{m,1}-\tilde{v}_{m,1})^{\top}g_{i,0}-\frac{1}{2}\tau\gamma(v_{i,1}-\tilde{v}_{i,1})^{\top}g_{m,0}\\=&\delta_{m,i}+\frac{1}{4}\tau^2\gamma^2\left[9g_{m,1}^{\top}(g_{i,1}-g_{i,0})+6(g_{m,1}-g_{m,0})^{\top}g_{i,0}-3g_{m,0}^{\top}(g_{i,1}-g_{i,0})\right]\\&-\frac{1}{2}\tau\gamma(v_{m,1}-\tilde{v}_{m,1})^{\top}g_{i,0}-\frac{1}{2}\tau\gamma(v_{i,1}-\tilde{v}_{i,1})^{\top}g_{m,0}.\end{aligned}$$

A similar derivation to (3.11) yields

$$\|g_{m,1}-g_{m,0}\|\le Q\tau,\quad \|g_{i,1}-g_{i,0}\|\le Q\tau. \tag{4.6}$$

By (4.2), we apply (4.6) to obtain

$$\begin{aligned}&\left|\tilde{v}_{m,2}^{\top}\tilde{v}_{i,2}-\delta_{m,i}\right|\\\le&\frac{1}{4}\tau^2\gamma^2\left[9\|g_{m,1}^{\top}\|\|g_{i,1}-g_{i,0}\|+6\|g_{m,1}-g_{m,0}\|\|g_{i,0}\|+3\|g_{m,0}\|\|g_{i,1}-g_{i,0}\|\right]\\&+\frac{1}{2}\tau\gamma\|v_{m,1}-\tilde{v}_{m,1}\|\|g_{i,0}\|+\frac{1}{2}\tau\gamma\|v_{i,1}-\tilde{v}_{i,1}\|\|g_{m,0}\|\le Q\tau^3.\end{aligned}$$

Then an application of Lemma 4.1 leads to

$$\|v_{i,2}-\tilde{v}_{i,2}\|\le Q_2\tau^3,\quad 1\le i\le k, \tag{4.7}$$

for some $Q_2>0$. Similarly to the proof of $n=2$, we can show that

$$\begin{aligned}\left|\tilde{v}_{m,q}^{\top}\tilde{v}_{i,q}-\delta_{m,i}\right|\le&\frac{1}{4}\tau^2\gamma^2\left[9\|g_{m,q-1}\|\|g_{i,q-1}-g_{i,q-2}\|+6\|g_{m,q-1}-g_{m,q-2}\|\|g_{i,q-2}\|\right.\\&+3\|g_{m,q-2}\|\|g_{i,q-1}-g_{i,q-2}\|+\|g_{m,q-2}-g_{m,q-3}\|\|g_{i,q-2}\|\\&\left.+\|g_{i,q-2}-g_{i,q-3}\|\|g_{m,q-2}\|\right]\\&+\frac{1}{2}\tau\gamma\|v_{m,q-1}-\tilde{v}_{m,q-1}\|\|g_{i,q-2}\|+\frac{1}{2}\tau\gamma\|v_{i,q-1}-\tilde{v}_{i,q-1}\|\|g_{m,q-2}\|\end{aligned} \tag{4.8}$$

for $3\le q\le N_T$. Similar to the derivation of (3.18), we obtain

$$\begin{aligned}&\|g_{i,q-1}-g_{i,q-2}\|\le Q\tau+Q\|v_{i,q-1}-\tilde{v}_{i,q-1}\|,\\&\|g_{m,q-1}-g_{m,q-2}\|\le Q\tau+Q\|v_{m,q-1}-\tilde{v}_{m,q-1}\|.\end{aligned} \tag{4.9}$$

Substituting (4.9) into (4.8) yields

$$\begin{aligned}|\tilde{v}_{m,q}^{\top}\tilde{v}_{i,q}-\delta_{m,i}|\le&\bar{Q}\tau^2\|\tilde{v}_{m,q-2}-v_{m,q-2}\|+\bar{Q}\tau^2\|\tilde{v}_{i,q-2}-v_{i,q-2}\|\\&+\bar{Q}\tau\|\tilde{v}_{m,q-1}-v_{m,q-1}\|+\bar{Q}\tau\|\tilde{v}_{i,q-1}-v_{i,q-1}\|+\bar{Q}\tau^3,\end{aligned} \tag{4.10}$$

for some $\bar{Q}>0$. Setting $q=3$ and invoking (4.2)-(4.7) we get $|\tilde{v}_{i,3}^{\top}\tilde{v}_{j,3}-\delta_{i,j}|\leq Q\tau^3$ such that an application of Lemma 4.1 leads to

$$\|v_{i,3}-\tilde{v}_{i,3}\|\leq Q_3\tau^3,\quad 1\leq i\leq k \tag{4.11}$$

for some $Q_3>0$. Let $\tilde{Q}>0$ be a constant such that $\tilde{Q}>\max\{Q_2,Q_3,g((1+\varepsilon)\bar{Q})\}$ for some $\varepsilon>0$, and we intend to use mathematical induction to show that

$$\|v_{i,p}-\tilde{v}_{i,p}\|\leq\tilde{Q}\tau^3,\quad 1\leq i\leq k \tag{4.12}$$

holds for $2\leq p\leq N_T$. By the definition of $\tilde{Q}$ and (4.7)-(4.11), (4.12) holds for $p=2,3$. Suppose (4.12) holds for $p=q-1$ and $p=q-2$ for some $4\leq q\leq N_T$. Then it follows from (4.10)-(4.11) that

$$\left|\tilde{v}_{m,q}^{\top}\tilde{v}_{i,q}-\delta_{m,i}\right|\leq\bar{Q}\tau^3(2\tilde{Q}\tau^2+2\tilde{Q}\tau+1).$$

For τ sufficiently small we have $2\tilde{Q}\tau^2+2\tilde{Q}\tau\leq\varepsilon$ such that $|\tilde{v}_{m,q}^{\top}\tilde{v}_{i,q}-\delta_{m,i}|\leq(\varepsilon+1)\bar{Q}\tau^3$. From Lemma 4.1, we have

$$\|v_{i,q}-\tilde{v}_{i,q}\|\leq g\big((\varepsilon+1)\bar{Q}\big)\tau^3\leq\tilde{Q}\tau^3$$

for $1\leq i\leq k$, which is exactly (4.12) with $p=q$. Thus, the proof is completed by induction. □

4.3 Error estimates

Define the errors

$$e_n^x:=x(t_n)-x_n,\quad e_{i,n}^v:=v_i(t_n)-v_{i,n},\quad 1\leq n\leq N_T,\quad 1\leq i\leq k.$$

We then perform the error estimates for the Adams explicit scheme (4.1) of k-SD (2.1) in the following theorem.

Theorem 4.1. *Suppose Assumption* 2.1 *holds. Then the following estimate holds for τ sufficiently small:*

$$\big\|e_n^x\big\|+\sum_{i=1}^{k}\big\|e_{i,n}^v\big\|\leq Q\tau^2,\quad 1\leq n\leq N_T.$$

Here Q depends on k,L,T and β but is independent from τ,n and N_T.

Proof. Similar to the derivations in Theorem 3.1 we could bound e_n^x in terms of $e_{i,n}^v$ for $1\leq i\leq k$ as follows:

$$\big\|e_n^x\big\|\leq Q\tau\sum_{m=1}^{n-1}\sum_{j=1}^{k}\big\|e_{j,m}^v\big\|+Q\tau^2,\quad 1\leq n\leq N_T. \tag{4.13}$$

Similar to the derivations in Theorem 3.1 we could obtain the recurrence relation for the error e_n^v for $1\le i\le k$ as follows:

$$\begin{aligned}\left\|e_{i,n}^v\right\| \le \left\|e_{i,n-1}^v\right\| + Q\tau\left\|e_{i,n-1}^v\right\| + Q\tau^2\sum_{m=1}^{n-1}\sum_{j=1}^{k}\left\|e_{j,m}^v\right\| + Q\tau\sum_{j=1}^{i-1}\left\|e_{j,n-1}^v\right\| \\ + Q\tau\left\|e_{i,n-2}^v\right\| + Q\tau^2\sum_{m=1}^{n-2}\sum_{j=1}^{k}\left\|e_{j,m}^v\right\| + Q\tau\sum_{j=1}^{i-1}\left\|e_{j,n-2}^v\right\| + Q\tau^3.\end{aligned}$$

Adding this equation from $i=1$ to k and denoting

$$E_n^v := \sum_{i=1}^{k}\left\|e_{i,n}^v\right\|, \quad 1\le n\le N_T,$$

yield an estimate in terms of E_n^v

$$E_n^v \le E_{n-1}^v + Q\tau E_{n-1}^v + Q\tau^2\sum_{m=1}^{n-1}E_m^v + Q\tau E_{n-2}^v + Q\tau^3.$$

Adding this equation from $n=2$ to n_* and using

$$\tau^2\sum_{n=1}^{n_*}\sum_{m=1}^{n-1}E_m^v = \tau^2\sum_{m=1}^{n_*-1}\sum_{n=m+1}^{n_*}E_m^v \le T\tau\sum_{m=1}^{n_*-1}E_m^v,$$

we get

$$E_{n_*}^v \le E_1^v + Q\tau\sum_{n=1}^{n_*-1}E_n^v + Q\tau^2.$$

From [41, inequality (4.15)], we have $E_1^v \le Q\tau^2$. Then an application of the discrete Gronwall inequality leads to

$$E_n^v \le Q\tau^2, \quad 2\le n\le N_T.$$

Substituting this estimate back into (4.13) yields an estimate of $\|e_n^x\|$, and we complete the proof. □

Remark 4.1. In this paper, the first time step is discretized using a straightforward explicit Euler scheme. While this approach mathematically preserves the second-order accuracy of the global error, it is generally not advisable due to the potential for introducing a significantly larger constant in the error estimate. A more accurate approach for the initial step is recommended, as this can minimize the error. Practically, a warm-up numerical method, such as using a smaller step size or employing a predictor-corrector technique, is advisable.

Remark 4.2. The paper only studies error estimation for the two-step Adams method. Error estimation for higher-order Adams methods can be proven similarly but requires higher smoothness assumptions on the energy function.

5 Numerical experiments

In this section, we perform numerical experiments to confirm the accuracy of Adams explicit schemes (3.3) and (4.1), as well as the path convergence of the numerical solutions to the actual search paths of HiSD. We denote the error by

$$\begin{aligned}
&\mathrm{Err}(x) := \max_{1\le n\le N_T} \|x(t_n)-x_n\|,\\
&\mathrm{Err}(v) := \max_{1\le n\le N_T} \|v(t_n)-v_n\|,\\
&\mathrm{Err}(\tilde{v}) := \max_{2\le n\le N_T} |\|\tilde{v}_n\|-1|,\\
&\mathrm{Err}(v_i) := \max_{1\le n\le N_T} \|v_i(t_n)-v_{i,n}\|, \quad 1\le i\le k,\\
&\mathrm{Err}(\tilde{v}_i) := \max_{2\le n\le N_T} |\|\tilde{v}_{i,n}\|-1|, \qquad 1\le i\le k
\end{aligned}$$

to test their convergence rates. Since there is no exact solution for HiSD, the numerical solution calculated by $\tau=2^{-20}$ is used as the reference solution, and for simplicity we set $\beta=\gamma=T=1$.

Example 5.1 (Accuracy Test under Minyaev-Quapp Surface)**.** We consider the saddle dynamics for the Minyaev-Quapp surface [26]

$$E(x_1,x_2)=\cos(2x_1)+\cos(2x_2)+0.57\cos(2x_1-2x_2).$$

We first compute its index-1 saddle point via the explicit Euler and the Adams explicit scheme (3.3) with the initial conditions $x(0)=(1,1)^\top$ and $v(0)=(0,1)^\top$. The results of the numerical experiments are displayed in Tables 1 and 2, respectively, and we can see that the errors of $\mathrm{Err}(x)$ and $\mathrm{Err}(v)$ in the explicit Euler scheme have only a first-order convergence rate, and the error of $\mathrm{Err}(\tilde{v})$ has a second-order convergence rate, whereas all three errors in the Adams explicit scheme have a convergence rate that is one order higher than that of the explicit Euler scheme. This agrees with the result of the proved theorem. Next, we compute its index-2 saddle point via scheme (4.1) with the initial conditions $x(0)=(1,1)^\top$, $v_1(0)=(0,1)^\top$ and $v_2(0)=(1,0)^\top$. In fact, the index-2 saddle point at this point is strictly a local maximum point. Numerical results are presented in Tables 3 and 4, which demonstrate the relevant accuracy of the Adams explicit scheme (4.1) as proved in Lemma 4.2 and Theorem 4.1. The CPU times of the explicit Euler scheme and the Adams explicit scheme at different time steps are shown in Table 5, and it can be seen that the two schemes spend about the same amount of CPU time, but the error of the Adams explicit scheme is much smaller than the error of explicit Euler scheme, so there is a greater advantage in using the Adams explicit scheme. The images of the search paths for the index-1 saddle point and index-2 saddle point are shown in Fig. 1.

Table 1: Convergence rates of Euler explicit scheme [41] in Example 5.1.

τ	Err(x)	Conv. rate	Err(v)	Conv. rate	Err($\tilde{v}$)	Conv. rate
1/32	2.20E-02		1.73E-02		2.54E-03	
1/64	1.03E-02	1.09	8.35E-03	1.05	6.34E-04	2.00
1/128	5.03E-03	1.04	4.11E-03	1.02	1.59E-04	2.00
1/256	2.48E-03	1.01	2.04E-03	1.01	3.97E-05	2.00
1/512	1.23E-03	1.01	1.02E-03	1.01	9.92E-06	2.00

Table 2: Convergence rates of Adams explicit scheme (3.3) in Example 5.1.

τ	Err(x)	Conv. rate	Err(v)	Conv. rate	Err($\tilde{v}$)	Conv. rate
1/32	9.22E-03		4.58E-03		9.34E-4	
1/64	2.33E-03	1.98	1.19E-03	1.94	1.13E-4	3.05
1/128	5.85E-04	2.00	3.03E-04	1.98	1.41E-5	3.00
1/256	1.46E-04	2.00	7.65E-05	1.99	1.77E-6	2.99
1/512	3.66E-05	2.00	1.92E-05	1.99	2.22E-7	3.00

Table 3: Convergence rates of Adams explicit scheme (4.1) in Example 5.1.

τ	Err(x)	Conv. rate	Err(v_1)	Conv. rate	Err(v_2)	Conv. rate
1/32	1.94E-03		2.10E-03		2.10E-03	
1/64	5.05E-04	1.94	5.40E-04	1.96	5.40E-04	1.96
1/128	1.28E-04	1.97	1.37E-04	1.98	1.37E-04	1.98
1/256	3.24E-05	1.99	3.45E-05	1.99	3.45E-05	1.99
1/512	8.13E-06	1.99	8.66E-06	1.99	8.66E-06	1.99

Table 4: Convergence rates of Adams explicit scheme (4.1) in Example 5.1.

τ	Err($\tilde{v}_1$)	Conv. rate	Err($\tilde{v}_2$)	Conv. rate
1/32	3.44E-04		3.44E-04	
1/64	4.30E-05	3.00	4.30E-05	3.00
1/128	5.41E-06	2.99	5.41E-06	2.99
1/256	6.78E-07	3.00	6.78E-07	3.00
1/512	8.48E-08	3.00	8.48E-08	3.00

Table 5: CPU time (s) of Euler and Adams explicit scheme in Example 5.1.

CPU time(s)	$\tau=2^{-16}$	$\tau=2^{-18}$	$\tau=2^{-20}$	$\tau=2^{-22}$
Euler	0.62	2.19	8.55	33.81
Adams	0.62	2.21	8.83	34.82

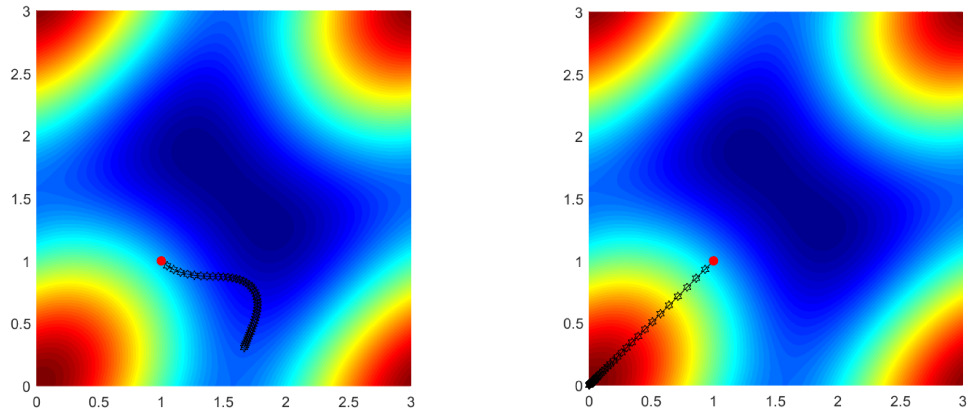

Figure 1: Numerical solution of $x(t)$ with $\tau=2^{-20}$ and terminal time $T=1$. (left) 1−SD, (Right) 2−SD.

Example 5.2 (Convergence of Dynamics under Modified Biggs EXP6 Functions)**.** Consider a six-dimensional Biggs EXP6 function [35]

$$B(\boldsymbol{x})=\sum_{i=1}^{6}\left(x_3\exp^{-t_ix_1}-x_4\exp^{-t_ix_2}+x_6\exp^{-t_ix_5}-y_i\right)^2,$$

where $t_i=i/10, y_i=\exp^{-t_i}-5\exp^{-10t_i}+3\exp^{-4t_i}$ and

$$B_k(\boldsymbol{x})=B(\boldsymbol{x})-\sum_{i=1}^{k}s_i\arctan^2(x_i-x_i^*)+\sum_{i=k+1}^{6}s_i\arctan^2(x_i-x_i^*).$$

By choosing $s=(4,8,16,8,4,2), x^*=(1,10,1,5,4,3)^\top$ becomes a k-saddle of $B_k(\boldsymbol{x})$ for $k=2,3,4,5$. An initial point $x(0)=(0,9,1,5,4,3)$ achieves convergence to x^*. The numerical results for the Adams explicit scheme (4.1) for $k=2,3,4,5$ are displayed in Table 6, where we can see that the Adams explicit scheme has second-order convergence speeds with respect to Err(x) and Err(v_1), third-order convergence speeds with respect to Err$(\tilde{v}_1)$, and numerical results with respect to Err(v_j) and Err$(\tilde{v}_j)$ for $j=2,3,4$ are omitted here due to similarities. The maximum allowable time steps for the explicit Euler method and the Adams explicit method are shown in Table 7. It can be observed that the maximum allowable time step for the explicit Euler method is approximately twice that of the Adams explicit method. The image of the error $\|x-x^*\|_2$ with respect to the number of iteration steps n about $k=4$ for different time steps is shown in Fig. 2. The figure shows that the coarser the step size, the faster the convergence to the saddle point, so we would prefer to use a coarse step size, but the first-order scheme at a coarse step size, because of the low precision, will result in an inaccurate path that may not converge to the specified saddle point. The second-order scheme solves this problem, and is still accurate at coarse step sizes.

Table 6: Convergence rates of Adams explicit scheme (4.1) in Example 5.2.

k	τ	Err(x)	Conv. rate	Err(v_1)	Conv. rate	Err($\tilde{v}_1$)	Conv. rate
2	1/64	1.05E-02		1.36E-03		4.67E-05	
	1/128	2.93E-03	1.84	3.32E-04	2.03	5.55E-06	3.07
	1/256	7.78E-04	1.92	8.29E-05	2.00	6.83E-07	3.02
	1/512	2.00E-04	1.96	2.08E-05	2.00	8.48E-08	3.01
	1/1024	5.09E-05	1.98	5.20E-06	2.00	1.06E-08	3.01
3	1/64	6.13E-03		8.83E-04		3.79E-06	
	1/128	1.61E-03	1.93	2.48E-04	1.83	3.23E-07	3.55
	1/256	4.13E-04	1.96	6.62E-05	1.91	3.99E-08	3.02
	1/512	1.05E-04	1.98	1.71E-05	1.95	4.95E-09	3.01
	1/1024	2.63E-05	1.99	4.35E-06	1.98	6.17E-10	3.01
4	1/64	6.13E-03		9.75E-04		4.28E-06	
	1/128	1.61E-03	1.93	2.73E-04	1.84	4.31E-07	3.31
	1/256	4.13E-04	1.96	7.25E-05	1.91	5.30E-08	3.02
	1/512	1.05E-04	1.98	1.87E-05	1.95	6.58E-09	3.01
	1/1024	2.63E-05	1.99	4.76E-06	1.98	8.19E-10	3.00
5	1/64	6.01E-03		9.73E-04		4.00E-06	
	1/128	1.58E-03	1.93	2.72E-04	1.84	3.90E-07	3.36
	1/256	4.05E-04	1.96	7.24E-05	1.91	4.76E-08	3.04
	1/512	1.03E-04	1.98	1.87E-05	1.95	5.90E-09	3.01
	1/1024	2.57E-05	1.99	4.73E-06	1.98	7.34E-10	3.01

Table 7: Maximum step size of Euler and Adams explicit scheme in Example 2.

Maximum step size	$k=2$	$k=3$	$k=4$	$k=5$
Euler	0.042	0.059	0.065	0.066
Adams	0.021	0.029	0.032	0.033

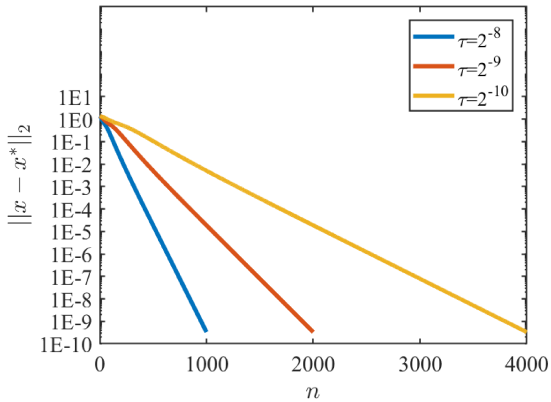

Figure 2: Plots of $\|x-x^*\|_2$ with respect to the iteration number for different time step τ.

6 Conclusions

In this paper, we first constructed the two-step Adams explicit scheme of HiSD. Then we develop new techniques to overcome the difficulties posed by strong nonlinearities and orthogonalisation procedures, proving that the Adams explicit scheme errors have second-order accuracy with respect to the time step. We rigorously demonstrate that the use of the explicit Euler scheme in the first step does not affect the accuracy of the later steps. The developed methods and results provide theoretical support for the accuracy of numerical calculations.

In contrast to the HiSD for gradient systems, the generalized HiSD [34] is used to compute any-index saddle points of the dynamical systems of non-gradient type, where the Hessian matrix in HiSD is replaced by a symmetrization of the Jacobian of the force. Similar to previous derivations, we could analyse the Adams explicit scheme of the generalized HiSD. For systems constrained by equalities, the constrained HiSD is developed in [31] to accommodate the constraints by employing the Riemannian gradients and Hessians. As the numerical treatments of Riemannian gradients and Hessians bring additional challenges, how to extend the proposed method to the constrained HiSD remains further investigations.

Compared to the explicit Euler method, the maximum step size of the Adams explicit method is reduced by about half. However, the corresponding computational accuracy is improved from first-order to second-order, which is beneficial for some high-dimensional problems [22]. High-order explicit schemes also include Runge-Kutta schemes. We know that the stability region of Runge-Kutta schemes expands with the increment of the order. The topic of constructing and analyzing higher-order explicit Runge-Kutta methods for higher-order saddle point dynamics will be found in [24].

Acknowledgments

This research was supported by the National Natural Science Foundation of China (Grant Nos. 12288101, 12225102, T2321001, 12301555), by the Taishan Scholars Program of Shandong Province (Grant No. tsqn202306083), by the National Key R&D Program of China (Grant No. 2023YFA1008903), and by the China Postdoctoral Science Foundation (Grant Nos. 2023M740103, GZB 20230028).

References

[1] J. Baker, *An algorithm for the location of transition states*, J. Comput. Chem., 7:385–395, 1986.
[2] J. Bofill, W. Quapp, and M. Caballero, *Locating transition states on potential energy surfaces by the gentlest ascent dynamics*, Chem. Phys. Lett., 583:203–208, 2013.
[3] H. Brunner, *Collocation Methods for Volterra Integral and Related Functional Differential Equations*, Cambridge University Press, 2004.

[4] E. Cancés, F. Legoll, M.-C. Marinica, K. Minoukadeh, and F. Willaime, *Some improvements of the activation-relaxation technique method for finding transition pathways on potential energy surfaces*, J. Chem. Phys., 130:114711, 2009.
[5] X. Cheng, L. Lin, W. E, P. Zhang, and A. Shi, *Nucleation of ordered phases in block copolymers*, Phys. Rev. Lett., 104:148301, 2010.
[6] J. Doye and D. Wales, *Saddle points and dynamics of Lennard-Jones clusters, solids, and supercooled liquids*, J. Chem. Phys., 116:3777–3788, 2002.
[7] Q. Du and L. Zhang, *A constrained string method and its numerical analysis*, Commun. Math. Sci., 7:1039–1051, 2009.
[8] W. E, C. Ma, and L. Wu, *A comparative analysis of optimization and generalization properties of two-layer neural network and random feature models under gradient descent dynamics*, Sci. China Math., 63:1235–1258, 2020.
[9] W. E, W. Ren, and E. Vanden-Eijnden, *String method for the study of rare events*, Phys. Rev. B, 66:052301, 2002.
[10] W. E and E. Vanden-Eijnden, *Transition-path theory and path-finding algorithms for the study of rare events*, Annu. Rev. Phys. Chem., 61:391–420, 2010.
[11] W. E and X. Zhou, *The gentlest ascent dynamics*, Nonlinearity, 24:1831–1842, 2011.
[12] W. Gao, J. Leng, and X. Zhou, *An iterative minimization formulation for saddle point search*, SIAM J. Numer. Anal., 53:1786–1805, 2015.
[13] N. Gould, C. Ortner, and D. Packwood, *A dimer-type saddle search algorithm with preconditioning and linesearch*, Math. Comp., 85:2939–2966, 2016.
[14] W. Grantham, *Gradient transformation trajectory following algorithms for determining stationary min-max saddle points*, in: Advances in Dynamic Game Theory. Annals of the International Society of Dynamic Games, Vol. 9, Birkhäuser, 639–657, 2007.
[15] Y. Han, Y. Hu, P. Zhang, and L. Zhang, *Transition pathways between defect patterns in confined nematic liquid crystals*, J. Comput. Phys., 396:1–11, 2019.
[16] Y. Han, J. Yin, Y. Hu, A. Majumdar, and L. Zhang, *Solution landscapes of the simplified Ericksen-Leslie model and its comparison with the reduced Landau-de Gennes model*, Proc. R. Soc. A, 477:20210458, 2021.
[17] G. Henkelman, G. Jóhannesson, and H. Jónsson, *Methods for finding saddle points and minimum energy paths*, in: Theoretical Methods in Condensed Phase Chemistry, Vol. 5, Springer, 269–302, 2002.
[18] G. Henkelman and H. Jónsson, *A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives*, J. Chem. Phys., 111:7010–7022, 1999.
[19] G. Henkelman and H. Jónsson, *Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points*, J. Chem. Phys., 113:9978–9985, 2000.
[20] A. Levitt and C. Ortner, *Convergence and cycling in walker-type saddle search algorithms*, SIAM J. Numer. Anal., 55:2204–2227, 2017.
[21] Y. Li and J. Zhou, *A minimax method for finding multiple critical points and its applications to semilinear PDEs*, SIAM J. Sci. Comput., 23:840–865, 2001.
[22] Y. Luo, X. Zheng, X. Cheng, and L. Zhang, *Convergence analysis of discrete high-index saddle dynamics*, SIAM J. Numer. Anal., 60:2731–2750, 2022.
[23] F. Mallamace, C. Corsaro, D. Mallamace, S. Vasi, C. Vasi, P. Baglioni, S. Buldyrev, S. Chen, and H. Stanley, *Energy landscape in protein folding and unfolding*, Proc. Natl. Acad. Sci. USA, 113:3159–3163, 2016.
[24] S. Miao, L. Zhang, P. Zhang, and X. Zheng, *Construction and analysis for orthonormalized Runge-Kutta schemes of high-index saddle dynamics*, Commun. Nonlinear Sci. Numer. Simul.,

145:108731, 2025.
[25] J. Milnor, *Morse Theory*, Princeton University Press, 1963.
[26] R. Minyaev, W. Quapp, G. Subramanian, P. Schleyer, and Y. Ho, *Internal conrotation and disrotation in $H_2BCH_2BH_2$ and diborylmethane* 1,3 *H exchange*, J. Comput. Chem., 18:1792–1803, 1997.
[27] J. Onuchic, Z. Luthey-Schulten, and P. Wolynes, *Theory of protein folding: The energy landscape perspective*, Annu. Rev. Phys. Chem., 48:545–600, 1997.
[28] D. Sheppard, R. Terrell, and G. Henkelman, *Optimization methods for finding minimum energy paths*, J. Chem. Phys., 128:134106, 2008.
[29] B. Shi, Y. Han, and L. Zhang, *Nematic liquid crystals in a rectangular confinement: Solution landscape, and bifurcation*, SIAM J. Appl. Math., 82:1808–1828, 2022.
[30] Z. Xu, Y. Han, J. Yin, B. Yu, Y. Nishiura, and L. Zhang, *Solution landscapes of the diblock copolymer-homopolymer model under two-dimensional confinement*, Phys. Rev. E, 104:014505, 2021.
[31] J. Yin, Z. Huang, and L. Zhang, *Constrained high-index saddle dynamics for the solution landscape with equality constraints*, J. Sci. Comput., 91:62, 2022.
[32] J. Yin, K. Jiang, A.-C. Shi, P. Zhang, and L. Zhang, *Transition pathways connecting crystals and quasicrystals*, Proc. Natl. Acad. Sci. USA, 118:e2106230118, 2021.
[33] J. Yin, Y. Wang, J. Chen, P. Zhang, and L. Zhang, *Construction of a pathway map on a complicated energy landscape*, Phys. Rev. Lett., 124:090601, 2020.
[34] J. Yin, B. Yu, and L. Zhang, *Searching the solution landscape by generalized high-index saddle dynamics*, Sci. China Math., 64:1801–1816, 2021.
[35] J. Yin, L. Zhang, and P. Zhang, *High-index optimization-based shrinking dimer method for finding high-index saddle points*, SIAM J. Sci. Comput., 41:A3576–A3595, 2019.
[36] J. Yin, L. Zhang, and P. Zhang, *Solution landscape of Onsager functional identifies non-axisymmetric critical points*, Phys. D, 430:133081, 2022.
[37] J. Zhang and Q. Du, *Shrinking dimer dynamics and its applications to saddle point search*, SIAM J. Numer. Anal., 50:1899–1921, 2012.
[38] L. Zhang, L. Chen, and Q. Du, *Morphology of critical nuclei in solid-state phase transformations*, Phys. Rev. Lett., 98:265703, 2007.
[39] L. Zhang, Q. Du, and Z. Zheng, *Optimization-based shrinking dimer method for finding transition states*, SIAM J. Sci. Comput., 38:A528–A544, 2016.
[40] L. Zhang, W. Ren, A. Samanta, and Q. Du, *Recent developments in computational modelling of nucleation in phase transformations*, Npj Comput. Mater., 2:16003, 2016.
[41] L. Zhang, P. Zhang, and X. Zheng, *Error estimates for Euler discretization of high-index saddle dynamics*, SIAM J. Numer. Anal., 60:2925–2944, 2022.
[42] L. Zhang, P. Zhang, and X. Zheng, *Mathematical and numerical analysis to shrinking-dimer saddle dynamics with local Lipschitz conditions*, CSIAM Trans. Appl. Math., 4:157–176, 2023.